

Teacher and Student Perceptions of the Criterion-Referenced Assessment of English in CSAT

Hyun-Ju Kim and Jong-Bai Hwang*

Kim, Hyun-Ju, & Hwang, Jong-Bai. (2022). Teacher and student perceptions of the criterion-referenced assessment of English in CSAT. *English Teaching*, 77(2), 87-107.

The present study examines Korean high school English teachers' and university students' perceptions about the criterion-referenced assessment of English in CSAT. A total of 400 participants (149 teachers and 251 students) answered an online questionnaire. Teachers showed more positive attitudes than students in their perceptions about the reduction of excessive competition among students, more communicative English classes, and the use of Korean and mathematics scores of the CSAT for the admission materials. Students, on the other hand, were more positive than teachers in their perceptions about the effects of the criterion-referenced assessment of English in CSAT on the reduction of private education and more high scorers in the CSAT. Regarding the alternatives to current criterion-referenced assessments of English in CSAT, teachers seem to emphasize the development of new question types for a valid criterion-referenced assessment system while students seem to focus on improving the communication skills through the English test in CSAT, not on the validity of the criterion-referenced assessment.

Key words: CSAT, criterion-referenced assessment of English, perceptions, validity

*First Author: Hyun-Ju Kim, Professor, Department of British and American Humanities, Dankook University

Corresponding Author: Jong-Bai Hwang, Professor, Department of English Education, Konkuk University; 120, Neungdong-ro, Gwangjin-gu, Seoul 05029, Korea; Email: jongbai@konkuk.ac.kr

Received 9 April 2022; Reviewed 26 April 2022; Accepted 20 June May 2022



© 2022 The Korea Association of Teachers of English (KATE)

This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0, which permits anyone to copy, redistribute, remix, transmit and adapt the work, provided the original work and source is appropriately cited.

1. INTRODUCTION

In Korean society, the meaning of the College Scholastic Ability Test (CSAT) is very important and has a significant social impact. It is natural that people in education have great interest in the change of the CSAT evaluation system but it has been pointed out that there is not much research related to the CSAT in the field of English education in Korea and a lack of research on evaluation policy (Kwon, 2015). However, since the introduction of absolute evaluation of English tests, there have been more research on investigating the effect of the transition of testing methods from norm-referenced to criterion-referenced assessments on students' English abilities and educational situations (Lee, 2019; Lee & Han., 2021; Lee, Jung, Yang, Lee, & Bae, 2020; Yun, Lee, Yang, Pae, & Lee, 2020). Since 1994, the CSAT used for college entrance exams has maintained the relative evaluation system until 2017, but constantly it has been reported that the relative evaluation method, the norm-referenced assessment, causes students' excessive competition and consequently the problem of private education. In addition, it has been pointed out that excessive learning has been induced due to the norm-referenced assessment and that there has been a limitation to improving English communication skills (Ministry of Education, 2014).

In addition, it has been reported that there have been more students who received the first level on the assessment of English in CSAT compared to the periods of the relative evaluation (Lee, 2019; Lee et al., 2020; Ryu, 2019). This phenomenon has led to the perception of English being a less important subject than other major ones such as the Korean language and mathematics for the students preparing for college entrance exams, which caused people to increasingly neglect the English subject. Thereby, the importance of English as subject seems to have been decreasing. In this respect, negative aspects of the criterion-referenced assessment of English in CSAT have been frequently pointed out (Lee, 2019). As such, discussions on side effects, controversies, and issues that appear after the introduction of the criterion-referenced assessment of English in CSAT are constantly being raised (Korean Council of English-Related Academic Organizations, 2018; Lee, 2019; Ryu & Jang, 2018; Ryu, 2019).

Moreover, it was revealed that instructors at universities were also expressing concerns about the lack of students' general English skills (Lee et al., 2020). There was a concern that the quality of college education is also decreasing due to the low levels of students' English and it was pointed out that it might be because of switching the evaluation system from relative to absolute. In addition, many on-site high school English teachers have pointed out the negative washback effect of the criterion-referenced assessment of English in CSAT since the number of English classes in high schools have been gradually decreasing compared to other major subjects such as Korean language and mathematics (Korean English Federation, 2018). Therefore, this study attempts to examine the overall perception

about implementing the criterion-referenced assessment of English in CSAT from high school English teachers and university students. Research on the students' perceptions of the criterion-referenced assessment of English in CSAT is a necessary task for validation of the assessment and will provide a next procedure of validation of the CSAT, which would increase the reliability and validity of the CSAT and could make further attempts on English education and the evaluation system. With the need for this research, the current researchers have raised specific questions as follows:

1. What are the perceptions of high school English teachers and university students on the validity of the criterion-referenced assessment of English in CSAT?
2. What are the ways of improving the implementation of the criterion-referenced assessment of English in CSAT?

2. LITERATURE REVIEW

2.1. Relative and Absolute Criterion of Evaluation

Evaluation is a judgement or decision-making process by comparing obtained assessment results with a criterion, and it can be defined as the systematic gathering of information for the purpose of making decisions (Linn, 1980; Turnbull, 1989; Weiss, 1972). Therefore, the assessment process is a prerequisite of the evaluation process and the evaluation is established with a certain criterion. In this regard, criterion is an important tool that can be used to determine the accuracy of an evaluation. There are two types of criteria, relative criterion and absolute criterion, and the assessment decisions are made based on these criteria. The term, 'criterion-referenced evaluation' is used to show the level of achievement that a student reached the achievement standards rather than the norms. The 'norm-referenced evaluation' is to show a student's achievement differences among students and produce a dependable rank order of students across a continuum of achievement from high to low achievers (Stiggins, 1994). Therefore, the criterion of the norm-referenced assessment is the one obtained after a test is done, and it is dependent on the results within the group. In other words, norm-referenced assessment has a purpose of distributing the assessment results within a certain group and revealing differences between students while criterion-referenced assessment reveals a student's success as a ratio of performance for which success can be considered sufficient in a particular context. For example, in a situation like college admissions where a rank order or a quota is necessary, norm-referenced assessments are appropriate (Omstein, 1993).

Criterion-referenced assessments are used to reflect the progress of development of individual students' abilities in a particular context. Therefore, the criterion should be a pre-determined absolute value with cut-off points, and needs to be applied for everyone fairly. However, the criterion and the cut-off points should be carefully determined by educators, content specialists, assessment professionals, politicians, and other decision-makers, especially in large-scale testing situations (McClarty, Way, Porter, Beimers, & Miles, 2013). In summary, the differences between norm-referenced and criterion-referenced assessments are the purpose of the assessments, design of assessment, method of score interpretation, and presentation.

2.2. English Assessment in CSAT

Based on the results of the recently studied criterion-referenced assessment of English in CSAT (Lee, 2019; Lee et al., 2020; Ryu, 2019), it was revealed that the ratio of the first level in the English subject of the CSAT was 10.03% in 2018, 5.30% in 2019, 7.43% in 2020, and 12.66% in 2021. Those results, which were conducted for the last five years, showed that the first level ratio fluctuated unstably compared to the ones (4.62% in 2015, 4.62% in 2016, and 4.49% in 2017) during the period of the norm-referenced assessment. The fluctuation has been large every year during the criterion-referenced assessment period compared to the one during the norm-referenced assessment period. As a result, the new evaluation method, criterion-referenced assessment has been pointed out as the evaluation method which is hard for us to guess how well the students have to do on the English test to get the first level in CSAT. Although it was expected that the new evaluation system, the absolute evaluation, might be a bit weak in discriminating students' English abilities, the criterion-referenced assessment of English in CSAT has been conducted since 2018 to solve the problems presented above (Ministry of Education, 2014). This transition from norm-referenced to criterion-referenced assessment can bring about great changes to the test users such as students, English teachers, administrators for college admission, and parents.

Some research conducted before the new assessment system was applied stated that there were negative perceptions and attitudes toward the new policy, believing that it would cause greater problems in private education and aggregate the burden of both students and parents. In addition, communicative language teaching would not be able to be promoted as long as the question items of the English test in CSAT English do not change (Choi & Choi, 2016; Lee & Kim, 2018). Additionally, teachers pointed out that their overcrowded classes and lack of teacher training would make it hard to change their classrooms to more communication-centered ones. Regarding students' perceptions, much research found that there were positive perceptions and attitudes toward

criterion-referenced assessments with the expectation that the burden of learning English would decrease while there were negative perceptions of changes in English teaching and learning methods and improvement of English communication skills (Kang, 2016; Lee, 2017; Lee, 2018; Yi, 2016).

3. RESEARCH METHOD

A mixed methods design was used to examine the perceptions of the criterion-referenced assessment of English in CSAT and the impacts of the assessment on English education in general. The questionnaire data mean scores were computed to estimate the general perception of the criterion-referenced assessment of English in CSAT. The written responses of open-ended questions in the questionnaire were analyzed using content analysis in order to identify the main themes of the improvements of current criterion-referenced assessment of English in CSAT.

3.1. Participants

The total number of participants in this study was 400 (149 high school English teachers and 251 university students). Table 1 below presents the information of teachers' backgrounds. The 149 teachers were all full-time English teachers and had teaching experience for at least five years, and almost half of them (47.65%) had 10 to 19 years of English teaching experience. Of the 149 teachers, 70 teachers (46.98%) were working in Seoul or metropolitan cities, and 61 of them (40.94%) were at mid-sized or small cities. Regarding the type of school, 126 teachers (84.56%) were working at general high schools.

TABLE 1
Teachers' Backgrounds

Background	Number	Proportion (%)
Age	Under 30	6
	30 to 39	55
	40 to 49	58
	Over 50	30
Teaching experience	Under 5 years	13
	5 to 9 years	28
	10 to 19 years	71
	Over 20 years	37
School location	Metropolitan areas	70
	Mid-sized/Small cities	61
	Rural areas	18
School type	General	126
	Special Purpose	8
	Autonomous Private	3
	Autonomous Public	5
	Vocational	3
	Arts	0
	Others	4

Regarding the students, 251 participants (160 freshmen, 38 sophomores, 30 juniors, and 23 seniors) from different cities in South Korea participated in this study, majoring in different areas such as humanities, education, business, engineering, etc. Table 2 below presents the information of students' backgrounds. For the high school region of the participants, 83 (33.1%) were in Seoul or metropolitan cities, 130 (51.8%) in mid-sized cities, and 31 (12.4%) in rural areas studied in high school. For the gender distribution of the participants, there were 155 (61.8%) female and 96 (38.2%) male students. When looking at the university entrance year, most of the participants (229, 91.2%) were students who experienced the criterion-referenced assessment of English in CSAT. Most of the participants (212, 84.5%) graduated from general high schools and some of them graduated from self-reliant private high schools (12, 4.8%) or from self-reliant public high schools (6, 2.4%). Regarding the type of school, 212 students (84.5%) were studying at general high schools.

TABLE 2
Students' Backgrounds

Background	Number	Proportion (%)
Experience of C-R assessment	Yes	229
	No	22
Grade	Freshman	160
	Sophomore	38
	Junior	30
	Senior	23
Major	Humanities/Social Studies	138
	Natural Science/Engineering	49
	Education	45
	Others	19
School location	Metropolitan areas	83
	Mid-sized/Small cities	130
	Rural areas	31
School type	General	212
	Special Purpose	4
	Autonomous Private	12
	Autonomous Public	6
	Vocational	5
	Arts	5
Others	7	

Note. C-R Assessment refers to the criterion-referenced assessment.

3.2. Data Collection

The questionnaire was designed to investigate teacher and student perceptions of the English assessment in CSAT by the current researchers and conducted by Google Forms (See Appendix I and II). The questionnaire had three parts: (1) participants' background information such as gender, regions, assessment experience, high school types, and others. (2) their perceptions of the current criterion-referenced assessment of English in CSAT, and (3) their thoughts on the alternatives and improvements of current English assessment in CSAT. In the first part, participants' responses were analyzed using SPSS, and descriptive statistics were mainly used for analysis. The second part of the questionnaire contained items from which participants were asked to provide general perceptions of the English assessment in CSAT. The participants were asked to respond to 5 Likert scale items. Item responses were scored such that the higher the scale score, the more positive the participants' perceptions were about the use of criterion-referenced assessment of English in CSAT. Items were scored 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree. The open-ended question was to uncover participants' thoughts on the ways of improving current criterion-referenced assessment of English in CSAT. The

question was as follows: 'Feel free to describe your thoughts on the current criterion-referenced assessment of English in CSAT and about how to improve the English assessment for college entrance.' The answers to that question were used as supportive data to strengthen the quantitative data results.

3.3. Data Analysis

The online questionnaire results were tabulated so that a higher mean would indicate a stronger disagreement of the statement. For example, if a student selected 5 ('strongly agree') as his or her response to the statements such as 'the absolute evaluation of the current CSAT English means the criterion-referenced assessment,' and 'the criterion-referenced assessment of English in CSAT contributes to the transition of English classes to communication-centered,' the student's response would indicate a strong agreement of the statements regarding the concept and the positive influence of the criterion-referenced assessment of English in CSAT. The students' responses were then processed using descriptive statistical analysis techniques to calculate the percentage of responses based on the provided Likert scale. All data were normally distributed, therefore an independent t-test was conducted to identify the perception differences between the teachers and the students. The latter part of the questionnaire was to obtain opinions for the alternatives and improvements of the current criterion-referenced assessment and the participants' responses were analyzed using a Mann-Whitney U-test. The qualitative data from the open-ended question in the questionnaire were analyzed using content analysis. The researchers identified the frequent comments on the responses related to the topic of this research and examined the emerging patterns.

4. RESULTS

4.1. Perceptions of the Criterion-Referenced Assessment of English in CSAT

Regarding how much the participants understand the criterion-referenced assessment, 79.19 percent of teachers answered they knew it well, 18.82 percent said they knew it, and 2.68 percent said they did not know much about it. On the other hand, only 17.79 percent of students answered they knew it well, 47.8 percent said they knew it, and 34.3 percent said they did not know much about it. These results indicate that most teachers were aware of the criterion-referenced assessment while most students were not. That might be

because compared to the teachers, the students were not familiar with the term ‘criterion-referenced assessment’ itself.

Regarding the meaningfulness of the criterion-referenced assessment, 28.9 percent of teachers and 37.1 percent of students agreed with the current English assessment in CSAT being used as a criterion-referenced assessment while 30.9 percent of teachers and 14.4 percent of students disagreed with it. These results indicate that teachers and students perceive criterion-referenced assessment very differently. Regarding the degree of competition, only 16.1 percent of teachers agreed that excessive competition for the English subject was greatly reduced by the criterion-referenced assessment of CSAT English. On the other hand, most of students (71.7%) agreed with it. In addition, 51.1 percent of teachers and 12.4 percent of students disagreed with it, indicating that students mostly had positive perceptions of the use of the criterion-referenced assessment of English in CSAT in that they became less competitive in English due to the evaluation system, although teachers still considered that CSAT is still leading excessive competition in English.

Regarding private education, 24.1 percent of teachers and 62.2 percent of students disagreed with the statement that private education has decreased after using the criterion-referenced assessment of English in CSAT. These results show that unlike what teachers believe, students consider that private education still takes place a lot and has not been influenced by the new English evaluation system in CSAT, which is not originally expected by the Ministry of Education. Regarding the university admission process, the perceptions of teachers and students were shown very differently. Compared to students (57.3%), much fewer teachers (12.8%) disagreed with the importance of using the Korean language and mathematics scores on the CSAT as admission screening materials instead of English. Regarding the large number of high scorers in English, the perceptions between teachers and students were also very different. Almost half of the teachers (53%) disagreed that it is desirable for students to receive the first level more easily and that there are a large number of first level students. On the other hand, only 32.7 percent of the students disagreed with them.

Finally, as for the communication-centered teaching and learning, it was shown that compared to the students (68.1%), much fewer teachers (7.4%) did not think that the criterion-referenced assessment of English in CSAT influenced teaching and learning methods. 56.4 percent of teachers and 14.4 percent of the students agreed with the statement, ‘The criterion-referenced assessment of English in CSAT contributes to the transition of English class to communication,’ indicating that more students than the teachers considered that their English classes have not been changed to the communication-centered one.

4.2. Perception Differences between Teachers and Students

The means of the responses on each item regarding the perceptions of the current criterion-referenced assessment of English in CSAT were analyzed using a *t*-test to test significant differences between teachers and students in their perceptions. Table 3 below shows the perception differences between teachers and students. Significant differences were found in all 6 items. The results are found in Table 3. Overall, students were a bit more positive than teachers toward using the criterion-referenced assessment of English in CSAT because excessive competition has decreased and because there are more possibilities for them to get high scores. However, they did not consider that it contributed to reducing private education and to communication-centered classes. On the other hand, teachers were somewhat more positive than students because they considered that the criterion-referenced assessment of English in CSAT contributed to reducing private education to some extent and leading to communication-centered classes, which is totally opposite to the students' perceptions.

TABLE 3
Perception Differences between Teachers and Students

<i>Item</i>	<i>Group</i>	<i>M</i>	<i>SD</i>	<i>df</i>	<i>Mean difference</i>	<i>t</i>	<i>Sig.</i>
Meaningfulness	T	3.01	.983	398	-.280	-2.938	.003
	S	3.29	.884				
Excessive competition	T	2.62	.969	398	-1.236	-11.651	.000
	S	3.86	1.059				
Private education	T	3.07	.871	398	.711	6.063	.000
	S	2.36	1.265				
Admission process	T	3.43	.808	398	.963	9.281	.000
	S	2.47	1.104				
High scorers	T	2.68	1.128	398	-.351	-2.992	.003
	S	3.04	1.140				
C-C class	T	3.50	.732	398	1.361	12.683	.000
	S	2.14	1.182				

Note. C-C Class refers to the communication-centered class.

4.3. Alternatives and Improvements for Criterion-Referenced Assessment of English in CSAT

In the questionnaire, the teachers and students were asked to respond to five items pertaining to the alternatives to current criterion-referenced assessments of English in CSAT. As shown in Table 4, in general, it was revealed that the teachers and students considered that English in CSAT should be conducted by criterion-referenced assessment as in its original purpose. In particular, the teachers (38.7%) answered the most to the

alternative, 'Not only English, but also Korean and mathematics should be evaluated by the criterion-referenced assessment.' The students (23.2%) answered the second to the alternative, 'English in CSAT should be as it was,' which implies that they did not want to have frequent changes in the evaluation system.

TABLE 4

Alternatives to Current Criterion-Referenced Assessment of English in CSAT

Alternatives	N	P	PoC	T	S
1. Not only English, but also Korean and mathematics should be evaluated by the criterion-referenced assessment.	113	23.4%	28.2%	70(38.7%)	43(14.2%)
2. English in CSAT should return to the norm-referenced assessment as before.	97	20.1%	24.3%	41(22.7%)	56(18.5%)
3. English in CSAT should be conducted by criterion-referenced assessment as in its original purpose.	123	25.5%	30.8%	48(26.5%)	75(24.8%)
4. English in CSAT should be as it is.	84	17.4%	21.0%	14(7.7%)	70(23.2%)
5. Others	66	13.7%	16.5%	8(4.4%)	58(19.2%)
Total	483	100.0%	120.8%	181(100%)	302(100%)

Note. PoC = percent of case; T = teachers; S = students

To verify if there were significant differences on the perceptions of the alternatives to current criterion-referenced assessments of English in CSAT between teachers and students, a Mann-Whitney U test was carried out. Since p-values in three items of alternatives were .000, we could conclude that there were significant differences in the perceptions of the three alternatives between the teachers and the students, but since the last one was 'others,' only two would be analyzed in this section. As shown in Table 5, there were significant differences between the teachers and the students in 'Not only English, but also Korean and mathematics should be evaluated by the criterion-referenced assessment' and in 'English in CSAT should be as it is.' These results indicate that teachers are more concerned with fairness of assessment while students are more concerned with stability of assessment.

TABLE 5
Perception Differences of Alternatives between Teachers and Students

	Alt. 1	Alt. 2	Alt. 3	Alt. 4	Others
Mann-Whitney U	13118.000	17726.000	18263.000	15241.500	15382.500
Wilcoxon W	44744.000	49352.000	49889.000	26416.500	26557.500
Z	-6.403	-1.173	-.489	-4.385	-4.615
P	.000	.241	.625	.000	.000

Note. Alt. 1 = 'Not only English, but also Korean and mathematics should be evaluated by the criterion-referenced assessment.'; Alt. 2 = 'English in CSAT should return to the norm-referenced assessment as before.'; Alt. 3 = 'English in CSAT should be conducted by criterion-referenced assessment as in its original purpose.'; Alt. 4 = 'English in CSAT should be as it is.'

Finally, there were five items asking about improvements of current criterion-referenced assessment of English in CSAT. As shown in Table 6, it was found that the teachers and students mostly considered that the difficulty level of the English in CSAT should be appropriated like the results of the Korean language and mathematics. Specifically, the teachers (35.4%) answered the most to the item for improvement, 'New question types suitable for the criterion-referenced assessment of English in CSAT should be developed and implemented.' The students (25.4%) answered the second most to the item, 'Speaking and writing sections should be included in CSAT English,' which implies that they wanted to improve their speaking and writing skills by preparing for the exam.

TABLE 6
Improvements of Current Criterion-Referenced Assessment of English in CSAT

Improvements	N	P	PoC	T	S
1. Speaking and writing sections should be included in CSAT English.	117	21.5%	28.3%	33(15.6%)	84(25.4%)
2. Testing methods should be improved (e.g. item banking system).	92	16.9%	23.0%	38(17.9%)	54(16.3%)
3. New question types suitable for the criterion-referenced assessment of English in CSAT should be developed and implemented.	140	25.8%	35.0%	75(35.4%)	65(19.6%)
4. The difficulty level of the English in CSAT should be appropriated according to Korean language and mathematics.	142	26.2%	35.5%	54(25.5%)	88(26.6%)
5. Others	52	9.6%	13.0%	12(5.7%)	40(12.1%)
Total	543	100.0%	135.8%	212(100%)	331(100%)

A Mann-Whitney U test was carried out again to verify the significant differences between teachers and students in the perceptions of improvements of current criterion-referenced assessments of English in CSAT. As shown in Table 7, it was revealed that there were significant differences in two items of improvements between the teachers and the students, which were in 'Speaking and writing sections should be included in

CSAT English,’ and ‘New question types suitable for the criterion-referenced assessment of English in CSAT should be developed and implemented.’ These results indicate that teachers consider developing new question types as more important to make the assessment valid than any other ways while students consider that English speaking and writing skills should also be assessed to improve the criterion-referenced assessment of English in CSAT.

TABLE 7
Perception Differences of the Improvements Between Teachers and Students

	Imp. 1	Imp. 2	Imp. 3	Imp. 4	Others
Mann-Whitney U	16583.000	17953.500	14129.500	18478.500	17225.500
Wilcoxon W	27758.000	48579.500	45755.500	50104.500	28400.500
Z	-2.403	-.915	-4.948	-.239	-2.264
P	.016	.360	.000	.811	.024

Note. Imp. 1 = ‘Speaking and writing sections should be included in CSAT English.’; Imp. 2 = ‘Testing methods should be improved (e.g. item banking system).’; Imp. 3 = ‘New question types suitable for the criterion-referenced assessment of English in CSAT should be developed and implemented.’; Imp. 4 = ‘The difficulty level of the English in CSAT should be appropriated according to Korean language and mathematics.’

4.4. Comments on Criterion-Referenced Assessment of English in CSAT

A number of issues related to current criterion-referenced assessment of English in CSAT were revealed by the data collected from the open-ended question in the student questionnaire. First, the data revealed that the English assessment in CSAT needs to include English productive skills like speaking and writing. The following excerpts seem to indicate that students’ concern is on authenticity, which means that the test items need to be closely related to real life.

Our English assessment in CSAT seems to have focused too much on reading. I think we need a new test method to improve our speaking skills.

I think an evaluation of speaking and writing abilities should be implemented in the criterion-referenced assessment of English in CSAT.

I think we should evaluate speaking, writing, reading and listening, I mean all, in our criterion-referenced assessment of English in CSAT.

Maybe the current test method can be best to use only as a material for the college entrance exam. However, since we do not learn speaking and writing skills at school, it’s hard for us to use English in a real life and I think it would be better if we study

speaking and writing along with reading and listening and take the test like TOEFL and use the scores for college admissions.

At universities, I believe, English speaking and writing skills are really important, but there is no section for those skills. I think CSAT is for a study at universities; then, English assessment in CSAT needs to evaluate students' general English ability including all four skills.

Second, it was suggested that the criterion-referenced assessment of English in CSAT should be fair and valid. The following excerpts are some of the fairness and validity issues raised by the participants:

Honestly, I'm not sure it would be good if we have the criterion-referenced assessment of English only in CSAT. I think every subject should have the same evaluation method. Why only English? It doesn't make sense at all. I mean criterion-referenced assessment for every subject or norm-referenced assessment for every subject.

After we have the criterion-referenced assessment of English in CSAT, we're making a lot more efforts in Korean and mathematics since it says the discriminating power in English is not very high. I doubt that it really matches the original purpose of criterion-referenced assessment.

I think we should go back to use the norm-referenced assessment. If we can't, we should have criterion-referenced assessment for all subjects in CSAT. I don't think the criterion-referenced assessment of English only doesn't make sense at all. It just feels like the importance of English is continuously falling down.

Thirdly, it was suggested that various types of items should be developed and included in the criterion-referenced assessment of English in CSAT. The following excerpts are some of the examples showing the suggestion regarding test items:

Changing test items. We have had consistently the same type of test items such as 'finding topics,' 'finding titles,' 'filling blanks,' 'ordering sentences/paragraphs,' 'finding summaries,' etc. for a long time. I doubt these types of items really assess the achievement of English learners. I don't think spending 12 years on learning English to solve these kinds of fixed test items.

I think there should be much more different types of test items in the criterion-referenced assessment of English in CSAT so that students' English ability can be easily discriminated.

I don't think it is necessary to make the test difficult to improve the level of discrimination. Rather, we need more various types of items assessing real communicative competence, not just vocabulary, grammar, and reading.

I understand it is not realistic that we have short-answer questions in CSAT, but I don't think all multiple-choice questions are right. Maybe various types of reading passage? For example, newspaper or magazine articles, I mean, some kinds of practical reading passages like the reading passages in TOEIC.

Finally, but most significantly, it was suggested that the most important part of CSAT is to maintain the same level of difficulty. Especially, in order for a valid criterion-referenced assessment of English in CSAT to be made, the participants said it would be necessary to maintain the difficulty level. The following excerpts are some of the examples stating the importance of a difficulty level:

I think the main value of using criterion-referenced assessments is not to determine whether a student's English level is relatively high or low compared to other students but to be closely related to improving English ability and alleviating competition among students. So, for this purpose, I think the test should adjust the level of difficulty appropriately.

We often find that foreigners get frustrated when they take the English test in CSAT because of some of difficult test items. I think it's a big problem. I mean if the test is difficult for English speakers, how difficult would it be for us?

If we would continue to use the criterion-referenced assessment of English in CSAT, I think it would be desirable to set the difficulty level to about four percent, or maybe it would be better to return it to the norm-referenced assessment.

5. DISCUSSION AND CONCLUSION

The present study attempted to examine high school English teachers' and university students' overall perceptions about the criterion-referenced assessment of English in CSAT, and to compare the two groups' perceptions. It also explored their perceptions about alternatives and ways of improvement of the test. It is surprising that for all the statements regarding the concept and the influence of the criterion-referenced assessment of English in CSAT, teachers and students showed significantly different responses. English teachers showed more positive attitudes than students in their perceptions about the reduction of excessive competition among students, more communicative English classes, and the use of Korean and mathematics scores of the CSAT for the admission screening materials. Students, on the other hand, were more positive than teachers in their perceptions about the effects of the criterion-referenced assessment of English in CSAT on the reduction of private education and more high scorers in the CSAT. These results of this research are in line with previous studies on the absolute evaluation system of CSAT (Lee, 2019; Ryu, 2019; Yi, 2016).

The large difference between students and teachers on their perceptions of the criterion-referenced assessment of English in CSAT may be due to the difference in their different positions on the CSAT. English teachers seem to be well aware of the purposes of the new evaluation system of English in CSAT, that is, the reduction of the excessive competition among students and the negative effects of high school English education only for the CSAT. Students, on the other hand, might have experienced the reduction of excessive competition for university entrance due to the implementation of the new CSAT evaluation system. It is also possible that the proportion of private tutoring or attending private institutes for English has actually decreased. In addition, the increase of the number of the high scorers of English in CSAT may have resulted in the students' positive attitudes toward the new criterion-referenced assessment of English in CSAT. However, since the proportion of the first level of English in CSAT has been fluctuated, further research is necessary. For instance, for the CSAT in 2022, the proportion of the first level of English was very low (6.25%) compared to the one in the previous year (12.66%). This fluctuation is the biggest problem and challenge in the stabilization of the criterion-referenced assessment.

The results of the survey about the alternatives and improvements of current English assessment in CSAT also showed significant differences between English teachers and students, even though both of them seem to recognize the problems of the new criterion-reference assessment of English in CSAT to a considerable extent. Regarding the alternatives to current criterion-referenced assessments of English in CSAT, significant differences are found in 'Not only English, but also Korean and mathematics should be

evaluated by the criterion-referenced assessment' (Ts: 38.7% vs. Ss: 14.2%), and in 'English in CSAT should be as it is' (Ts: 7.7% vs. Ss: 23.2%). These results, as mentioned before, may indicate that teachers are more concerned with fairness of assessment while students are more concerned with stability of assessment. These differences seem to be related to the results regarding their perception of improvements of current criterion-referenced assessments of English in CSAT, in which significant differences were found in two items of improvements, which were in 'New question types suitable for the criterion-referenced assessment of English in CSAT should be developed and implemented' (Ts: 35.4% vs. Ss: 19.6%), and in 'Speaking and writing sections should be included in CSAT English' (Ts: 15.6% vs. Ss: 25.4%). Teachers seem to emphasize the development of new question types for a valid and adequate criterion-referenced assessment system, while students seem to focus on improving the communication skills through the English test in CSAT, not on the validity of the criterion-referenced assessment.

The results of the present study are expected to provide several pedagogical implications for English language teaching and learning in high school classrooms in Korea. First of all, the data from the open-ended question not only revealed the participants' concern about the English test in CSAT, but also implied a desirable direction of English education in Korea. The students pointed out one of the critical problems of the English test in CSAT, which focuses too much on reading skills, and insisted the test be changed to assess students' productive skills such as speaking and writing. Such changes in CSAT are directly related to changes in ways of learning and teaching English in high school classrooms. The students seem to be well aware of one of the biggest problems of English language teaching and learning in Korea, which focuses too much on reading skills and is not related to the use of English in real life. In other words, many students want to learn speaking and writing skills as well as listening and reading at school, and we need to learn and teach English in a more communicative way, with more opportunities to use English which is closely related to real life.

This study has some limitations in data collection from university students, not from high school students, and in a further study, it would be better to include current high school students' perspectives and compare them with the results from the previous research (Lee, 2019; Yi, 2016) and the current one. To be more specific about the limitation of data collection and analysis, this research analyzed the responses on the open-ended questions from students only and when discussing the results, a few main points must be considered carefully. As this simple research on the perception of the criterion-referenced assessment of English in CSAT highlights, teachers and students approach the assessment system from a very different aspect. Lee and Han (2021) keenly observed that many have been overwhelmed with the issues and challenges of the current

criterion-referenced assessment of English in CSAT. It is not the issue for teachers and students only, but for everyone in many educational sectors. Under the situation that there is no country in Asia where the criterion-referenced assessment of English in a large-scale test is conducted except for South Korea, now is the right time to discuss the future with this assessment system and design reliable and valid pedagogical approaches and the entire evaluation system. Current researchers believe that with sufficient research in various contexts and by solving the problems that have been raised, the criterion-referenced assessment of English will no longer be an issue of criticism or something to get rid of in the CSAT evaluation system. Instead, it will be a key to solve many issues in English education in South Korea.

Applicable levels: Secondary, tertiary

REFERENCES

- Choi, S. J., & Choi, J. (2016). Perceptions of English teachers on the new criterion-referenced assessment in the CSAT. *The Journal of English Language & Literature*, 58(4), 371-401.
- Kang, H. (2016). *A study of learners' perception on absolute grading system in the English section of CSAT*. Unpublished master's thesis, Sookmyung Women's University, Seoul, Korea.
- Korean Council of English-Related Academic Organizations. (2018). *The crisis of English education in Korea: English education division and national competitiveness*. Paper presented at the Korean Council of English-Related Academic Organizations Joint Symposium. Seoul, Korea.
- Kwon, O. (2015). A history of policies regarding the English section of Korea's College Scholastic Ability Test. *English Teaching*, 70(5), 3-34.
- Lee, J. A. (2019). *Perceptions of high school students and teachers on criterion-referenced evaluation in the English section of CSAT*. Unpublished master's thesis, University of Seoul, Seoul, Korea.
- Lee, J. Y. (2018). *A study about how the application of criterion-referenced evaluation in CSAT affects English studies*. Unpublished master's thesis, Yonsei University, Seoul, Korea

- Lee, K., & Kim, J. (2018). A study on high school English teachers' perceptions and responses to English education policy. *Korean Education Inquiry*, 36(4), 71-95.
- Lee, Y. (2017. 12.13). *Changing demand for English and Korean realities (in the era of the 4th Industrial Revolution)*. The Annual International Conference of the English Language and Literature Association of Korea. Seoul National University, Seoul, Korea.
- Lee, Y., & Han, S. (2021). KCSAT English grades and college scholastic abilities: A survey of college students' perceptions. *The 21st Century Association of English Language and Literature*, 34(4), 419-452.
- Lee, Y., Jung, H., Yang, Y., Lee, J., & Pae, H. (2020). Investigating changing trends in college English education before and after the introduction of the jeoldaepyeongga system to the KCSAT English section. *The Journal of English Language & Literature*, 66(4), 808-846.
- Lee, Y. S. (2019). Implementing a policy of absolute grading for CSAT English in Korea: A case of politically embedded test. *English Teaching* 74(2), 3-25.
- Linn, R. L. (1980). Issues of validity for criterion-referenced measures. *Applied Psychological Measurement*, 4(4), 547-561.
- McClarty, K. L., Way, W. D., Porter, A. C., Beimers, J. N., & Miles, J. A. (2013). Evidence-based standard setting: Establishing a validity framework for cut scores. *Education Researcher*, 42(2), 78-88.
- Ministry of Education. (2014). *Introduction of absolute evaluation of English in the College Scholastic Ability Test*. Retrieved December 21, 2019 from <https://www.moe.go.kr/sn3hcv/doc.html?fn=5dcb5c742949a25dee2cc580e77ff6d3&rs=/upload/synap/202206/>
- Omstein, A. C. (1993). Norm-referenced and criterion-referenced tests: An overview. *NASSP Bulletin*, 77(1), 28-39.
- Ryu, J., & Jang, I. (2018). A critical discourse analysis of the policy of criterion-referenced assessment in the CSAT English section: A case of instructors in the private sector. *The SNU Journal of Education Research*, 27(3), 1-28.
- Ryu, J. (2019). Issues and tasks in implementation of jeoldaepyeongga for CSAT. *NARS Analysis of Current Issues*, 50. 1-16.
- Stiggins, R. J. (1994). *Student-centered classroom assessment*. New York: Merrill.
- Turnbull, J. M. (1989). What is ... normative versus criterion-referenced assessment. *Medical Teacher*, 11(2), 145-150
- Weiss, C. H. (1972). *Evaluation research: Methods for assessing program effectiveness*. Englewood Cliffs, NJ: Prentice Hall.

- Yi, J. H. (2016). *A study of students' perceptions on criterion-referenced evaluation of CSAT (College Scholastic Ability Test) in 2018*. Unpublished master's thesis, Incheon National University, Incheon, Korea.
- Yun, D., Lee, Y., Yang, Y. Y., Pae, H., & Lee, J. (2020). Analysis of the transition of English ability of freshmen in two engineering colleges and the situation of English education after the adoption of the criterion-referenced system for the English section of the College Scholastic Ability Test. *The 21st Century Association of English Language and Literature*, 33(4), 129-155.

APPENDIX I

Part of questionnaire items on the criterion-referenced assessment of English in CSAT

Items	1	2	3	4	5
The absolute evaluation of the current English assessment in CSAT is used as the criterion-referenced assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The over-competition for English subjects has decreased a lot because of the criterion-referenced assessment of English in CSAT.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Private education in English has decreased a lot due to the criterion-referenced assessment of English in CSAT.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
It is desirable that universities use Korean and mathematics scores as admission materials rather than English ones for discrimination purpose.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
It is desirable that there are a number of high-scoring (the first level) students in English in CSAT.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The criterion-referenced assessment contributes to the transition of English classes to communication.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Note. 1 = Strongly disagree; 2 = Disagree; 3 = So so; 4 = Agree; 5 = Strongly agree

APPENDIX II

Questionnaire items on alternatives and improvements for the criterion-referenced assessment of English in CSAT

Items	Options
Currently, only English subject are conducted in the criterion-referenced assessment and Korean and mathematics are conducted in norm-referenced assessment. As a response to this phenomenon, you can suggest the following. Choose the one you think is most appropriate (Multiple	<input type="checkbox"/> Not only English, but also Korean and mathematics should be evaluated by the criterion-referenced assessment. <input type="checkbox"/> English in CSAT should return to the norm-referenced assessment as before. <input type="checkbox"/> English in CSAT should be conducted by criterion-referenced assessment as in its original purpose. <input type="checkbox"/> English in CSAT should be as it is.

responses possible).	<input type="checkbox"/> Others
Currently, it is pointed out that the English assessment in CSAT has many problems for college entrance exams. What do you think is the most necessary way to improve the English assessment in CSAT? (Multiple responses possible)	<input type="checkbox"/> Speaking and writing sections should be included in CSAT English. <input type="checkbox"/> Testing methods should be improved (e.g. item banking system). <input type="checkbox"/> New question types suitable for the criterion-referenced assessment of English in CSAT should be developed and implemented. <input type="checkbox"/> The difficulty level of the English in CSAT should be appropriated according to Korean language and mathematics. <input type="checkbox"/> Others
